AD-A088 240 LITTLE (ARTHUR D) INC CAMBRIDGE MA F/G 11/5

A REVIEW AND ANALYSIS OF FIBER COUNTING METHODS.(U)

JUN 80 N60921-78-C-A221

UNCLASSIFIED NSWC-TR-80-210 NL

1 OF 1

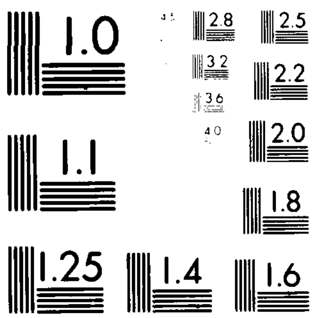
AD A088240

END

DATE FILMED

8-80

DTIC



MICROCOPY RESOLUTION TEST CHART

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(19) REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (18) NSWC/TR-80-210	2. GOVT ACCESSION NO. AD-A088 240	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) A REVIEW AND ANALYSIS OF FIBER COUNTING METHODS.		5. TYPE OF REPORT & PERIOD COVERED (9) Final rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Arthur D. Little, Inc. (12) 53		8. CONTRACT OR GRANT NUMBER(s) (15) N60921-78-C-A221
9. PERFORMING ORGANIZATION NAME AND ADDRESS Arthur D. Little, Inc. Acorn Park Cambridge, Massachusetts 02140		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS A510-5104/004-D/9W0652-0000 and Others
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Surface Weapons Center (F56) Dahlgren, Virginia 22448		12. REPORT DATE (11) June 1980
		13. NUMBER OF PAGES 56
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) (16) WA (17) WA 2.0000		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Fiber counting methods for use with carbon fibers are reviewed and analyzed. Recommendations are made for their use and comparison.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED 20??50 dlt

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

FOREWORD

The work described in this report was performed by Arthur D. Little, Inc., Cambridge, Massachusetts, under contract to the Naval Surface Weapons Center, Dahlgren, Virginia (Contract No. N60921-78-C-A221). The work was part of the carbon fiber hazard assessment program sponsored by NAVAIR-50 under AIRTASK A510-5102/004-F/9W0463-000, Work Unit A5203-01 and others. This report reviews the fiber counting methods used by the different groups working with this problem and recommends procedures to be used as standards for future counting efforts. It also provides factors to be applied if the results on different methods are compared.

This document has been reviewed and approved by J. H. Meyers, Navy Project Leader; C. E. Gallaher, Navy Project Manager of the Special Projects Branch; and L. J. Lysher, Head, Electromagnetic Effects Division.

Released by:

D. B. Colby

D. B. COLBY, Head
Electronics Systems Department

ACKNOWLEDGEMENT

This study was performed on behalf of the Electronics Systems Department of the U.S. Naval Surface Weapons Center by Mr. Gerald R. Larocque under the direction of Dr. Phani Raj.

The U.S. Navy's technical representative was Mr. J. H. Meyers whose support and cooperation are gratefully acknowledged. Also acknowledged are the many helpful comments of Mr. Joseph Diachenko and Mr. C. E. Gallaher of NSWC, Mr. John Threthway of DPG, Dr. Vernon Bell of NASA, and Dr. William Vickers of Mitre Corporation.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

ANALYSIS OF CARBON FIBER COUNTING PROCEDURES

1. INTRODUCTION

The high strength and low weight characteristics of carbon fiber composites make them extremely attractive materials for the fabrication of structural components for high performance aircraft. The electrical conductivity of the fibers, however, presents a potential threat to nearby electrical equipment in the event that fibers are released during an aviation accident. The U.S. Naval Surface Weapons Center (NSWC) has attempted to establish the important physical parameters governing such releases by conducting a series of small-scale tests at the Environmental Test Facility in Dahlgren, Virginia. These tests involve subjecting small samples of composite materials to propane fire, sometimes followed by explosive agitation, and collecting samples of the released fibers on sticky papers (6" x 9") placed on the chamber floor. Because of the extremely large number of fibers that are released, direct counting of the fibers on even a single sticky paper is infeasible, and instead a statistical technique has been employed to estimate the fibers released.

The statistical fiber counting procedures involve counting either actual number of fibers in a very small area or counting fiber intersections with parallel lines, again in a small area, and then extrapolating the result to the whole sticky paper area. This type of counting procedure necessarily involves scatter in the final result. It is therefore important to identify and quantify the errors inherent in the counting process. This report examines in detail the fiber counting procedure used in the Dahlgren tests (as implemented at the U.S. Army Dugway Proving Ground) and provides a comparison with counting techniques used by several other organizations.

The following sections detail the specific procedures being examined, present the theoretical basis of each procedure, and provide individual estimates for each source of uncertainty. The final result is that all of the available procedures provide sufficient accuracy for purposes of risk analysis and that the Dugway procedure is therefore preferred for its ease of implementation.

2. DESCRIPTION OF FIBER COUNTING PROCEDURES

Following a chamber test at the Dahlgren facility, sticky papers representing one quadrant of the test chamber are forwarded to the U.S. Army Dugway Proving Ground (DPG) in Dugway, Utah, for counting. In order to prevent fiber loss during transit, each paper is covered with a sheet of clear plastic and the sheets are enclosed in plastic envelopes. Upon receipt, the following procedure is used by DPG to estimate the number of fibers released.

2.1 Summary of the Dugway Counting Procedure

A. Estimation of mean fiber length

1. For each sheet to be counted a 1-square-inch area is randomly selected by positioning the sheet beneath a counting microscope.
2. A microscope is focused in the selected area and the lengths of 10 single fibers are measured. Single fibers are distinguished by diameter measurements using a graticule in the counting microscope.
3. This procedure is repeated for each of the approximately 30 sheets that are to be counted (1 quarter of the test chamber).
4. The resulting measurements are used to estimate the mean fiber length and length distribution for the tests.

B. Estimation of number of fibers on sheet

1. A set of parallel lines with spacing at least 20 percent larger than the mean fiber length calculated above is selected and placed on the sticky paper over an area of fairly uniform fiber distribution. The exact placement of the grid is at the discretion of the person counting the sample. A typical sticky paper record is shown in Figure 2.1.
2. The grid lines are scanned until 100 intersections have been counted (300 if the sample is very dense) and the following formula is used to estimate the number of fibers on the paper:

$$\hat{N} = \frac{\hat{I}\pi A}{2\ell\hat{L}} \qquad (2.1)$$

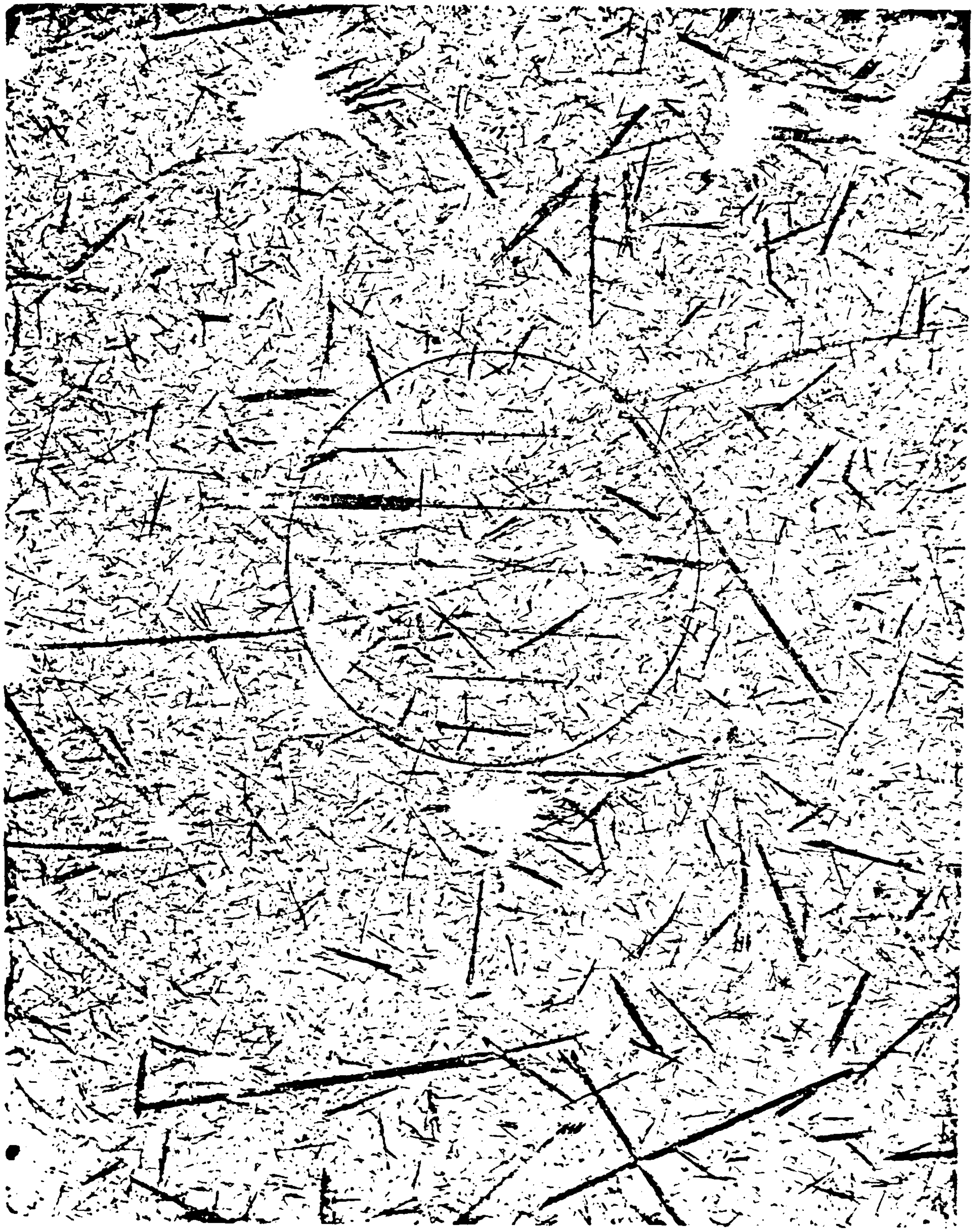


FIGURE 2.1

REPRESENTATIVE STICKY PAPER RECORD

where $\hat{N}$ = estimated number of fibers on sticky paper

$\hat{I}$ = number of intersections counted

A = area of sticky paper

L = length of line scanned to count I intersections

$\hat{\ell}$ = estimated average fiber length

The summary of results from this counting procedure* are recorded in table form, a sample of which is shown in Table 2.1.

The basis for the DPG counting procedure is the Buffon needle problem. This classical probability problem and some required extensions are discussed further in Appendix A.

2.2 Implications of the Extended Buffon Needle Problem

The analysis in Appendix A was developed to verify the accuracy of the DPG algorithm when applied to distributions containing fibers longer than the line spacing. It further serves to resolve the recent concerns by DPG that their procedure applies only to Gaussian distributions of fiber lengths and that for other distributions the procedure should be based on the median fiber length. This analysis shows that the current procedure is correct, regardless of distribution shape, if the mean fiber length is used. It will be shown later, however, that the shape of the fiber length distribution will affect the expected error of the estimate.

---

*It is important to note that two revisions to the counting procedure have been implemented during the past year. Originally, the set of parallel lines were of significant width relative to the fiber diameter. This situation would alter the probability of a fiber crossing a grid line and thus decrease the accuracy of the fiber count. A second change occurred in the identification of fibers longer than 1 mm. The person counting the sample is asked to select only those fibers longer than 1 mm but originally was not provided with reference marks of comparable length. Recently, 1-mm reference marks have been added to the grid lines (as shown in Figure 2.1). The implications of these modifications will be discussed later in this chapter.

TABLE 2.1

Dugway Proving Ground Test Data Sample

Test No. BT-232/X-178 Date: 7 Dec 1978

Total Singles in Room: 5051937

Mean Length of Singles: 4.1 mm Median: 3.4 mm

Standard Deviation: 2.6 mm

Length Category (mm)	% Frequency Distribution	% Cumulative Frequency
1-2	21.0	21.0
2-3	24.0	45.0
3-4	13.7	58.7
4-5	12.3	71.0
5-6	9.0	80.0
6-7	6.0	86.0
7-8	4.3	90.3
8-9	4.0	94.3
9-10	2.7	97.0
10-11	1.0	98.0
11-12	0.7	98.7
12-13	0.3	99.0
13-14	0.0	99.0
14-15	0.3	99.3
> 15	0.7	100.0

T-300/5208

HiTemp

30 x 30 x 0.64 cm

BT-232/X-178

2.3 Description of Alternative Counting Methods

A few other organizations have been involved in developing and using procedures for counting carbon fibers released from experiments conducted by these organizations. Amongst these are NASA (Langley), Scientific Services, Inc., (SSI) of California and TRW, Inc. Each organization has used variations of the same technique of counting, namely, counting a small region of a sample and scaling the result by the ratio of the total area of the room (or the sticky paper) to the counted sample area. Since these methods of statistical estimation of total fibers have a bearing on the method utilized by the DPG, they are discussed below.

2.3.1 The NASA-Langley Fiber Counting Procedure

1. A 3.5-cm square is cut from the lower right corner of each sample.
2. The opaque backing is removed, the transparent sticky portions are mounted on an aperture card and enlarged photographically (20x).
3. Fibers are counted by mounting the enlargement on a magnetic digitizer board and using a magnetic pen to identify the fiber endpoints to a computer. A computer program then computes and summarizes fiber lengths and count. For light depositions (less than 4×10^4 fibers/m^2), the entire enlargement is counted. For medium and heavy deposits, a 12.7-cm square is randomly selected near the center of the photograph and this area is counted.
4. The ratio of sample area to area counted is used to scale the count to the entire sticky paper.

2.3.2 The Scientific Systems, Inc., (SSI) Procedure

1. Four 8-mm x 8-mm regions are selected by visually identifying one area of light density, one of heavy density and two of intermediate density.
2. Each area is counted under a four-power stereo microscope and the fibers classified into 1-mm length intervals.

3. Total estimated fiber count is obtained by multiplying the counts by the ratio of total sampler collecting area to the area represented by the fiber counting regions.

2.3.3 The TRW Procedure

1. Ten random positions are selected on the sticky paper.
2. At each position, a 9-cm x 10-cm photograph is made at magnifications of 1, 3, 10, 30 and 100 power. The upper left corner of each enlargement is located at the randomly selected point.
3. The fibers on each enlargement are counted and scaled by the ratio of total sample area to area counted to produce an estimate of the total number of fibers on the sticky paper.

3. DISCUSSION OF THE TYPES OF STATISTICAL ERRORS IN THE FIBER COUNTING PROCEDURES

Since all of the counting procedures involve statistical estimation techniques, a number of errors are possible. The following list represents possible error sources associated with this procedure.

- Systematic error in estimation of length distribution
- Random error in estimation of length distribution
- Systematic error in estimation of fiber count
- Random error in estimation of fiber count
- Errors due to inaccurate or inconsistent applications of the technique
- Errors associated with extrapolation of counts
- Errors due to mistakes by personnel

Before discussing each error in detail, it is appropriate to briefly describe each type of error and the mathematical concepts used to quantify them.

A systematic error is a tendency to either overestimate or underestimate a particular parameter over many observations. In the present procedure, independent estimates will be developed for the length

distribution of the carbon fibers and for the total number of fibers released. It is important to know whether given many estimates the results will tend to be higher or lower than the real values. Analytically, this is represented by the concept of statistical bias which is in turn defined in terms of the mathematical expectation. Mathematical expectation is a fairly simple concept closely related to our everyday notion of average. In general if we seek to estimate something subject to sources of uncertainty, such as the number of fibers on a sticky paper, a different answer will be obtained each time it is estimated. If the probability of obtaining any particular answer can be calculated either theoretically or based on previous experience, it is possible to determine, a priori, what the average of many observations will be by multiplying the probabilities by the values and summing. This is the mathematical expectation, denoted $E\{\hat{x}\}$, where $\hat{x}$ is the variable being observed. The expectation is defined by

$$E\{\hat{x}\} = \sum_{i=1}^{n} \hat{x}_i p(\hat{x}_i) \tag{3.1}$$

where $\hat{x}_i$ are the possible values of $\hat{x}$ and $p(\hat{x}_i)$ is the probability of observing $\hat{x}_i$. If x can take on continuous values, its probabilities are expressed by the density function and the expectation is given by the following equation

$$E\{\hat{x}\} = \int_{-\infty}^{\infty} \hat{x} p(\hat{x}) dx \tag{3.2}$$

The concept of bias can now be best understood by noting that in general, if we seek to estimate some parameter, such as the total number of fibers on a sticky paper, there are many observations that may be used to approximate it. For example, we might observe values of the DPG estimator given in Equation 2.1, or we might count all the fibers in a 2-cm, or in a 4-cm circle. Any of these observations would be termed a statistic and an estimator for the total fiber count.

Figure 3.1 might indicate results of several (4) measurements using each of three different estimators. It can be seen that the first

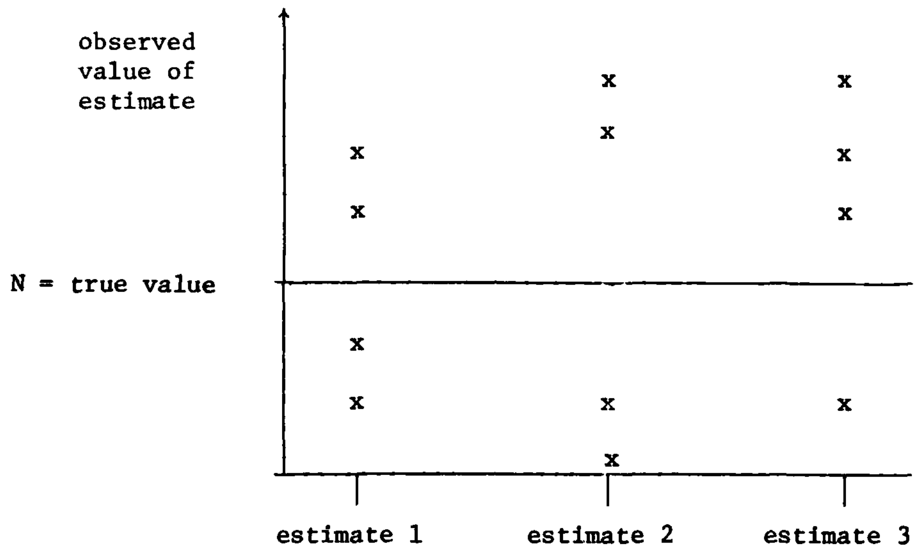


FIGURE 3.1

Illustration of Systematic Errors in Estimates

two estimators are "better" in that their average value would be equal to the true value whereas the third tends to overestimate the parameter. Estimators 1 and 2 are said to be "unbiased" while 3 is said to have a positive bias. Expressed mathematically, if $\hat{x}$ is used to estimate x

$\hat{x}$ is unbiased if $E\{\hat{x}\} = x$

$\hat{x}$ is positively biased if $E\{\hat{x}\} > x$

$\hat{x}$ is negatively biased if $E\{\hat{x}\} < x$

A random error represents a random fluctuation of an estimate around its average or expected value. It represents the degree of nearness we can expect from the true parameter value. This situation is illustrated in Figure 3.2.

Estimators 1 and 2 can both be seen to be unbiased or to have no systematic error since their average is the true value. However, estimator 1 is likely to yield a more reliable estimate of the parameter than is estimator 2 since the observations are more tightly grouped about the true value and, hence, one is less likely to observe a value far removed from the true value.

Estimator 3 is included to illustrate an interesting relation between types of errors. Although, in general, an unbiased estimate is preferred to a biased one, in this case, the bias is small compared to the random variation, and estimator 3 might still give more accurate results than estimator 2.

$$V\{x\} = \sum_{i=1}^{n} (\hat{x}_i - \overline{x})^2 \, p(\hat{x}_i) \tag{3.3}$$

or for continuous variables,

$$V\{x\} = \int_{-\infty}^{\infty} (\hat{x} - \overline{x})^2 p(\hat{x}) d\hat{x} \tag{3.4}$$

Since the variance measures the mean-square spread about the average value, it can be used to estimate the expected size of deviations. The Chebyshev Inequality provides a technique for this estimation. This inequality is stated as follows:

$$\Pr\{|\hat{x} - x| > \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} \tag{3.5}$$

where ε = arbitrary positive constant

and $\sigma^2 = V\{\hat{x}\}$ = variance of the estimate

x = true parameter value

$\hat{x}$ = estimated parameter value

This expression may be used to estimate confidence levels for the estimate $\hat{x}$ as follows. If α is a constant between 0 and 1, then the following relation is true with probability $1 - \alpha$.

$$|\hat{x} - x| \leq \sqrt{\frac{\sigma^2}{\alpha}} \tag{3.6}$$

or, with probability $1 - \alpha$, the fractional error in the estimate is less than:

$$e_f = \frac{1}{x}\sqrt{\frac{\sigma^2}{\alpha}} \tag{3.7}$$

For example, if $\alpha = 0.05$, then 95% of the parameter estimates will lie within $\pm\ e_f \cdot 100$ percent of the average value, e.g., if the variance (σ^2) of an estimate is 4 and the mean value is 200, then the 95% confidence interval is about ± 4% of the average value.

Finally, the notion of a consistent estimate should be introduced. This is also an indication of the expected level of random error. It represents the degree to which the observations will group around the mean as the sample size is increased. For example, if an estimate of the fiber release is based on counting a fixed area of the sticky paper, as the area becomes larger it is less and less likely that our estimate will be inaccurate. Mathematically, this is expressed by the following statement.

An estimate $\hat{x}$ is a consistent estimate of a parameter x if and only if

$$\lim_{m\to\infty} \Pr\{|\hat{x}(x_1 \ . \ . \ . \ x_m) - x| \geq \varepsilon\} = 0 \tag{3.8}$$

where $x_1 \ . \ . \ . \ x_m$ are observations,

x is a function of $x_1 \ . \ . \ . \ x_m$ used to estimate x

ε is an arbitrary positive number.

For purposes of the accidential release program, only a rough estimate of the number of released fibers is required. An accuracy to within a factor of 2 or 3 is sufficient for the evaluation of hazard associated with the release of carbon fiber. Consequently, estimation techniques need not be elaborately examined to determine the most precise technique. It is sufficient instead to examine the estimators for major statistical problems and base further recommendations on criteria such as ease of implementation and cost.

4. ESTIMATION OF ERRORS FOR THE DUGWAY PROCEDURE

One of the interesting properties of the released fiber is the distribution of fiber lengths since fiber length is related to associated hazard. In Step 4 of the DPG procedure, the fiber length distribution is estimated by counting the number of fibers in bins of 1 mm width and dividing by the total number of fibers counted. Since it is possible that a fiber could have any length, one is really interested in estimating a continuous length density function as illustrated in Figure 4.1. The best density function estimate based on the bins is determined by dividing the fraction of fibers in each bin by the width of the bin.

This procedure introduces a systematic error into the density function estimate that is a function of the bin size and the "true" continuous density function. This point is demonstrated by the following mathematics:

Let $\hat{p}(x)$ denote the fraction of fibers counted within a length bin of width Δx centered on length x. Thus, if $p(\ell)$ is the true length density function, the expectation of $\hat{p}(x)$ is given by the following formula:

$$E\ \{\hat{p}(x)\} = \frac{1}{\Delta x} \int_{x - \Delta x/2}^{x + \Delta x/2} p(\ell)\ d\ell \qquad (4.1)$$

Assuming that $p\ (\ell)$ is three times continuously differentiable,

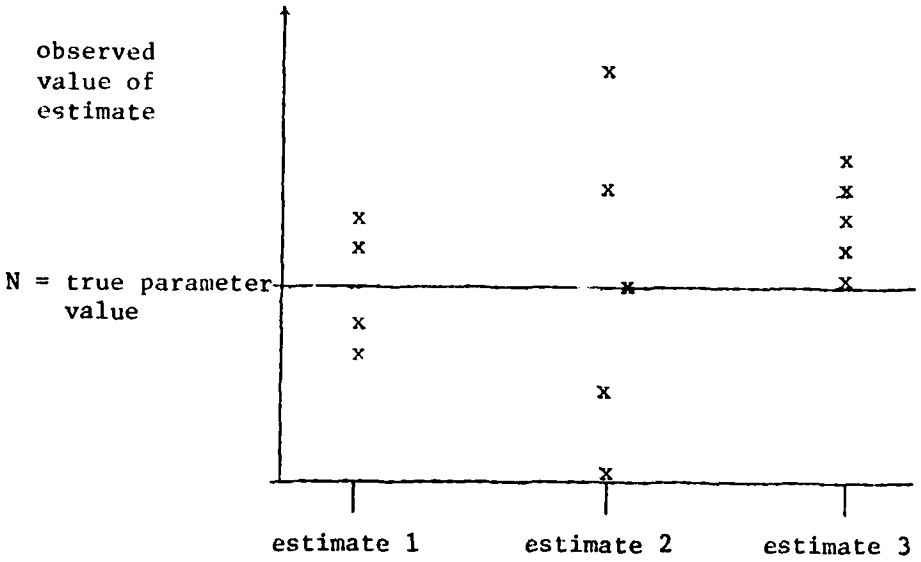


FIGURE 3.2

Illustration of Random Error in Estimates

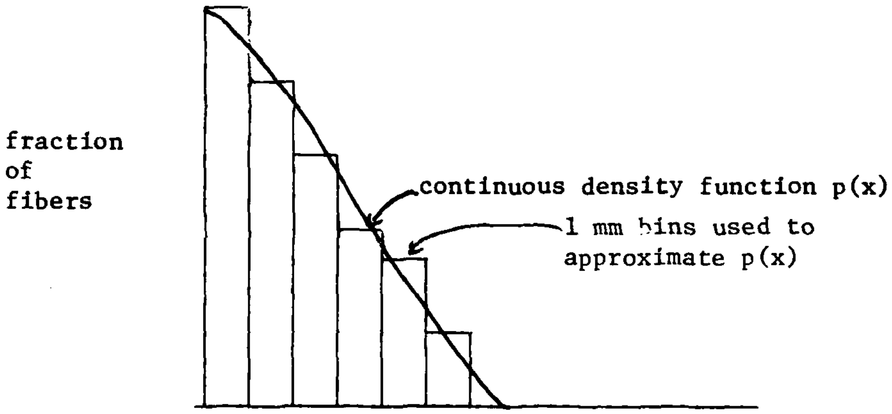


FIGURE 4.1

Estimation of Fiber Length Distribution

$p(\ell)$ may be expressed as a Taylor series with remainder:

$$p(\ell) = p(x) + (\ell-x)\frac{dp(x)}{dx} + \frac{(\ell-x)^2}{2}\frac{d^2p(x)}{dx^2} + \frac{1}{2}\int_x^{\ell}(x-t)^2\frac{d^3p(t)}{dt^3}\,dt \qquad (4.2)$$

Thus,

$$E\{\hat{p}(x)\} = p(x) + \frac{\Delta x^2}{24}\frac{d^2p(x)}{dx^2} + \frac{1}{2\Delta x}\int_{x-\Delta x/2}^{x+\Delta x/2}\int_x^{\ell}(x-t)^2\frac{d^3p(t)}{dt^3}\,dt\,d\ell \qquad (4.3)$$

Hence to within second order terms,

$$E\{\hat{p}(x)\} = p(x) + \frac{\Delta x^2}{24}\frac{d^2p(x)}{dx^2} \qquad (4.4)$$

Thus, $\hat{p}(x)$ is a biased estimate of the true density function $p(x)$. This bias, however, decreases with the square of the bin size and thus, for sufficiently small bins will be insignificant. Since $p(x)$ is an unknown function, it is not possible to accurately calculate this bias. Based on the experimental results, however, it has been seen that for fibers longer than 1 mm an exponential distribution will, generally, fit the data to within about 10%.* Although other distributions, such as the incomplete Gamma provide a better fit, the exponential is simple enough to allow analytic calculations and is sufficiently accurate for the present analysis. It thus seems appropriate to roughly estimate this error term based on an exponential distribution having a mean value of about 3 to 4 mm. The resulting calculations indicate an error of less than 5 percent, which cannot be considered a significant error source.

---

*The exponential distribution is discussed in Appendix B for readers who are unfamiliar with it.

4.1 Random Error in Estimation of Length Distribution

If the true length distribution p(x) of the released fibers is known, it is possible to estimate the random error or variance of the estimator $\hat{p}(x)$ as described in the previous section. Given p(x), the probability that a randomly chosen fiber will have a length falling between $x-\frac{\Delta x}{2}$ and $x+\frac{\Delta x}{2}$ (i.e., a bin of width Δx centered on x) is

$$P = \int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} p(x)\,dx \qquad (4.5)$$

Thus, if N fibers are selected and measured, the expected number n of fibers in this interval is given by:

$$E\{n\} = NP \qquad (4.6)$$

and the variance of n is given by:

$$V\{n\} = NP(1-P) \qquad (4.7)$$

But, since the estimate $\hat{p}(x)$ is defined by $\hat{p}(x) = \frac{n}{N\Delta x}$, the variance of $\hat{p}(x)$ is given by

$$V\{\hat{p}(x)\} = \frac{1}{N^2(\Delta x)^2} \quad V(n)$$

$$= \frac{1}{N^2(\Delta x)^2} \quad NP(1-P)$$

$$= \frac{P(1-P)}{N(\Delta x)^2} \qquad (4.8)$$

and thus decreases inversely with the number of fibers selected. In order to estimate the effect, p(x) will again be assumed exponential with a mean value of 3 mm. It then follows that

$$P = \int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} p(x)dx = e^{-0.33x}\left\{e^{\frac{0.33\Delta x}{2}} - e^{\frac{-0.33\Delta x}{2}}\right\} \qquad (4.9)$$

and for 1 mm length bins,

$$P \simeq 0.33e^{-0.33x} \qquad (4.10)$$

so the variance at x = 1 mm, for counting 300 fibers is given by

$$V\{\hat{p}(x)\} = \frac{0.24(1-0.24)}{N}$$

$$= \frac{0.18}{N} = \frac{0.18}{300} = 6.08 \times 10^{-4} \tag{4.11}$$

and the expectation is

$$E\{\hat{p}(1)\} = \frac{N \cdot 0.24}{N\Delta x} = 0.24 \tag{4.12}$$

So that the 95% confidence interval for $\hat{p}(1)$ is approximately 45% of its expectation.

4.2 Systematic Error in the Estimation of Fiber Count

It should be noted that there are two obvious approaches to the application of the estimator of Equation 2.1. .

1. Intersections may be counted for fixed length of line
2. The length of line measured after counting a specified number of intersections.

The second approach is currently followed by DPG, but both will be examined to determine if an improvement may be obtained. In general, the errors associated with fiber count estimation will be dealt with in the section comparing results of alternative methods. The following analytic expressions are included, however, for the qualitative insights they yield.

4.2.1 Error for Fixed Length Method

Under this interpretation, the results of Appendix A and the assumption of uniform spatial distribution may be used to rewrite Equation 2.1 as

$$E\{\hat{N}\} = \frac{\pi A}{2L} \cdot \frac{2\bar{\ell}N}{\pi D} \cdot \frac{L}{L_T} \cdot E\{\frac{1}{\ell}\} \tag{4.13}$$

$$= N\,\bar{\ell}\,E\{\frac{1}{\ell}\}$$

where L_T is the total line length to cover the sample so that N is unbiased if and only if

$$E\{\frac{1}{\hat{\ell}}\} = \frac{1}{\bar{\ell}} \quad . \tag{4.14}$$

This condition, however, is dependent upon the fiber length distribution and will generally result in a positive bias or a tendency to overestimate because Jensen's inequality implies that

$$E\{\frac{1}{\hat{\ell}}\} \geq \frac{1}{E\{\hat{\ell}\}} = \frac{1}{\bar{\ell}} \quad . \tag{4.15}$$

This bias is not in itself a serious shortcoming because the above estimate is consistent and based on a relatively large number of fibers, i.e., since many fibers are used to compute $\hat{\ell}$, the probability of $\frac{1}{\hat{\ell}}$ being very different from $\frac{1}{\bar{\ell}}$ is small.

The numerical estimation of this bias clearly depends on knowing the distribution of the average length. Even then, the distribution must be simple enough to permit the calculation of $E\{\frac{1}{\hat{\ell}}\}$ in order to estimate the bias. Unfortunately this does not appear to be true. However, for a particular experiment, this bias is fixed since $\hat{\ell}$ is not estimated separately for each sheet. It is thus possible to derive some rough approximation for the range of the bias from experiment to experiment as follows.

Since $\hat{\ell}$ is the sample mean of 300 observations, it will have a standard deviation smaller than the overall population by a factor of $1/\sqrt{300} = 5.8 \times 10^{-2}$. Since a typical test might have a mean fiber length of 3 mm and a standard deviation of 2 mm, $\hat{\ell}$ should have a mean of 3 mm and a standard deviation of $2/\sqrt{300} = 0.12$.

It thus follows from Equation 3.7 that at least 95% of the time, the estimate of $\hat{\ell}$ will be accurate to within 18%. Consequently, although the bias is unknown for any test, this effect can be expected to introduce an error of less than +22% or -15% for any test. In comparison to other error sources such as sampling errors and random errors associated with

the counting process, this is an unimportant contribution to the overall inaccuracy.

4.2.2 Error for Fixed Intersection Method

If, however, the line length is measured for a specified number of intersections, the expected value of $\hat{N}$ becomes

$$E\{\hat{N}\} = \frac{\pi AI}{2} E\{\frac{1}{\hat{\ell}}\} E\{\frac{1}{\hat{L}}\} \tag{4.16}$$

The above comments on the bias introduced by the term $\frac{1}{\hat{\ell}}$ are still applicable. However, additional bias will now be introduced unless

$$E\{\frac{1}{\hat{L}}\} = \frac{2N\bar{\ell}}{\pi AI} . \tag{4.17}$$

Thus, since

$$E\{\hat{L}\} = \frac{\pi AI}{2N\bar{\ell}} \tag{4.18}$$

it follows that

$$E\{\frac{1}{\hat{L}}\} \geq \frac{1}{E\{\hat{L}\}} = \frac{2N\bar{\ell}}{\pi AI} . \tag{4.19}$$

This application of the estimator will in general increase the positive bias of the estimate except for special distributions. The result is that if the counting process is repeated and the estimates averaged, the result will tend to differ from the true number of fibers. However, due to the consistency property, if relatively large numbers of intersections are counted and large numbers of fibers used to estimate the average fiber length, the result will tend to be near the true value of N.

Unfortunately, the distribution of the observed line length, $\hat{L}$, is a quite complex analytic expression and no empiric data are available to estimate it. Thus, the magnitude of this source of error cannot be numerically estimated by analytic means. The results of the

above sections are significant in that they demonstrate qualitatively that the method of counting intersections over a fixed line length provides a reduction in error over the use of the fixed intersection procedure.

From an operational standpoint, however, this application of the estimator of Equation 2.1 is not very different from the procedure described above if L is chosen as the expected length to count 100 intersections. It therefore appears reasonable on an intuitive basis to expect the difference in the errors to be small and use the above estimates to roughly approximate the error associated with this procedure.

4.3 Random Error in the Estimates of Fiber Count

As discussed above, this type of error is measured in terms of the variance of the estimate. For the alternative of counting intersections on a fixed length of line, the following analytic expression may be developed (Appendix B)

$$V\{\hat{N}\} = \left(\sum_{i=1}^{N} \frac{2\ell_i}{\pi D} \frac{A_s}{A} \left\{ 1 - \frac{2\ell_i A_s}{\pi D A} \right\} + \frac{A_s}{A} \sum_{i = N_1+1}^{N} \frac{4}{\pi} \sqrt{\left(\frac{\ell_i}{D}\right)^2 - 1} - \cos^{-1}\left(\frac{D}{\ell_i}\right) \right) \left(\frac{\pi \hat{A}}{2\bar{\ell}L} \right)^2 \quad (4.20)$$

where

$\hat{N}$ = estimated number of fibers on sticky paper

ℓ_i = length of i^{th} fiber

N_1 = number of fibers of length less than D

N = total number of fibers

D = line spacing

$\bar{\ell}$ = mean length of all fibers on the sticky paper

A_s = area counted (covered by line length L)

A = area of sticky paper

L = length of line counted

This expression is strongly dependent on the fiber length distribution. In order to provide an estimate of the approximate magnitude of this error, the data from test BT-230 were used as an example. This shows that the error due to this random variation should be less than a factor of 2 at the 95% confidence level. It should be cautioned, however, that since the fiber length distribution varies from test to test, the accuracy of this counting algorithm will differ with each test.

The probability distribution associated with counting a fixed number of intersections is quite complex, and it does not appear feasible to develop explicit analytic expressions for the variance. From an operational standpoint, however, the problem is very similar to counting a fixed line length whose length was chosen to be approximately that required to count 100 intersections. Although it cannot be rigorously demonstrated, it seems reasonable to expect the random errors for the two methods to be similar.

4.4 Errors due to Incorrect or Inconsistent Application of Techniques

Two significant sources of error not inherent to the DPG counting methodology have been noted. Initially, DPG counted using quite wide grid lines and no reference marks to aid in the identification of fibers of length 1 mm or greater. These two problems appear to introduce larger errors than those arising from the statistical estimation process. More significantly, the effects of these problems vary from experiment to experiment and can be expected to vary from counter to counter.

Table 4.1 shows the results of counting the same samples both with and without reference lines. As can be seen, the addition of reference marks has caused as much as a factor of 3 variation in the estimated fiber release and the factor is quite dependent on the density of the fiber deposit and probably on the person counting the sample. This implies a source of error which, although within the desired accuracy of counting, varies from experiment to experiment in a nonrandom fashion making comparison of results difficult.

No similar data are available to illustrate the effect of the wide grid lines.

In order to partially resolve this difficulty, the correction factors of Table 4.2 have been developed and checked with Dugway Proving Ground. It should be pointed out, however, that these factors

TABLE 4.1

Comparison of Fiber Counts With and Without 1 mm. Reference Marks in Dugway Counting Procedure

Test Number	Sampler No.	With Reference	Without Reference
BT-171	13863	1206	906
(Light	13875	1077	510
Deposit)	13882	876	477
	13884	857	317
BT-237	20803	25516	74202
(Medium	20805	14227	48464
Deposit)	20810	12718	44667
	20812	18782	38610
BT-230	19968	52071	120885
(Heavy	19970	47745	130995
Deposit)	19980	39151	83320
	20013	36883	96254

TABLE 4.2

Correction Factors for Absence of Reference Marks on Dugway Fiber Counting Procedure

Density of Fiber Deposition on the Sticky Paper	Range of Total Fiber Count	Correction Factor Applied to Results Obtained From Nonreference Mark Counts
Light	Less than 5×10^7	1.82
Medium	5×10^7 to 1.5×10^8	0.35
Heavy	Greater than 1.5×10^8	0.41

are very approximate and at best can only serve to partially eliminate the inconsistency in the counting technique.

4.5 Errors Associated with Extrapolation of Estimated Counts

The above analysis reflects the statistics of the intersection technique used to estimate the fiber deposit on a given sticky paper. A further error source is introduced by the extrapolation of each count to the 1-sq-meter section of the chamber that it represents. This extrapolation is clearly unbiased under the assumption of uniform fiber distribution. Its effect on the expected error is to multiply the variance of $\hat{N}$ by $\left(\frac{A}{A_s}\right)^2$, where A is the area represented and A_s is the area of the sticky paper. Consequently, the variance of the estimator is increased by approximately a factor of 900 due to this extrapolation. This means that the estimated fiber release for the 1 m^2 area is 30 times less accurate in terms of absolute reaction (see Equation 3.6) than the estimate for the individual sticky paper. When expressed as a percentage of total fiber count, however, the overall accuracy is uneffected by this extrapolation error.

5. STATISTICAL ERROR ESTIMATES FOR ALTERNATIVE COUNTING METHODS

Because of the similarity of the alternative fiber counting techniques, a single statistical analysis can be applied to the examination of each method. Each method involves counting a small area of the sample and scaling the result to represent the deposit on the entire sheet. Mathematically, this may be expressed by the following estimator:

$$\hat{N} = \frac{A}{A_s} n \qquad (5.1)$$

where $\hat{N}$ = estimated fiber count for the sticky paper

A = area of the entire sheet

A_s = area that was counted

n = number of fibers counted in the selected area

5.1 Systematic Errors in Alternative Counting Methods

Under the assumption that a fiber is equally likely to fall anywhere on the sticky paper, fibers are deposited on the area selected

for counting according to a binomial distribution with probability $p = \frac{A_s}{A}$.* It then follows that

$$E\{\hat{N}\} = \frac{A}{A_s} \; E\{n\} = \frac{A}{A_s} \; N \frac{A_s}{A} = N \tag{5.2}$$

where N is the true number of fibers on the sheet. Thus, the count in any randomly selected area is an unbiased estimate of the total fiber deposit. However, the established counting procedures do not always select the areas randomly, but sometimes base the area selection on the counter's perception of the fiber density. This practice will introduce biases that are impossible to quantify as the accuracy of the person's perception of fiber density is not known. The best that can be done is qualitative observations based on the assumption that it is possible to accurately locate the lowest, highest, and average density regions by visual perception.

The NASA-Langley estimates will clearly be unbiased since the region selected for counting is independent of the observed fiber density distribution. The TRW procedure is unbiased when the positions are randomly selected and probably introduces negligible bias when the counter attempts to locate a region of average density because his pattern recognition ability likely permits an accurate selection. The SSI procedure is positively biased, however, due to the application of equal weighting to regions of high- and low-fiber density. Since the expected number of fibers in the most dense region is further from the mean than the expected number of fibers in the least dense regions, a positive bias will be introduced assuming the counter accurately identifies these regions. Thus, the SSI procedure will tend to overestimate the number of fibers on the sheet. This will be seen in the following section where the results of a comparison counting are presented. The SSI results are consistently higher than those presented by other organizations.

---

*The binomial distribution is described in Appendix C for the benefit of readers unfamiliar with its properties.

5.2 Random Errors in Alternative Counting Procedures

All of the procedures will yield consistent estimates since they will tend to precise results as the size of the area counted approaches the size of the samplers.

The expected errors are represented by the variance of the estimate in Equation 5.1 and will decrease with the area counted since:

$$V\{\hat{N}\} = \frac{A^2}{A_s^2} V\{n\} = \frac{A^2}{A_s^2} N \frac{A_s}{A} \left(1 - \frac{A_s}{A}\right) \tag{5.3}$$

$$= N \left(\frac{A}{A_s} - 1 \right)$$

In order to interpret this result, some typical values must be assumed. For moderate to heavy releases, a sticky paper contains about 3×10^4 fibers and is approximately 15 x 23 cm. It is used to represent a 1-m^2 area of the test chamber. The NASA procedure counts a 3.5-cm-sq section of the sample, and the largest area counted by TRW is 8.75 x 10 cm.*

For moderate and heavy fiber densities, however, the NASA method does not count the entire 3.5 cm region, but rather a 12.7 cm square randomly located on a 20x enlargement. This implies that for the higher fiber densities, the region actually counted is only 0.64 cm. Using these values, Equation 5.3 may be used together with Equation 3.7 to develop some error estimates. The results of these calculations are presented in Table 5.1.

It is apparent that the errors associated with these algorithms are strongly dependent on the size of the area counted. The counting of a 4-cm^2 area, for example, would reduce the confidence interval for the NASA procedure to about ±29%. Although the variance associated with the current NASA procedure is large, it is still well within the factor of 3 accuracy required for the accidental release studies.

*In developing the error estimates for the TRW procedure, it has been assumed that this area was counted for one of the chosen points. The actual procedure used by TRW to combine results of the several photographs is, however, not well documented. In fact, the errors involved may be considerably larger than the estimate derived here as TRW may have, in fact, counted a smaller area.

TABLE 5.1

Estimated Extrapolation Variance and Error

	NASA	TRW
Variance for extrapolation to sample	855N	3N
Numerical estimate for sampler ($N = 3 \times 10^4$)	3×10^7	9×10^4
95% confidence for sampler	0.75	0.04

6. EMPIRIC COMPARISON OF COUNTING TECHNIQUES

In order to directly compare the results of the several counting techniques, four sticky papers were selected from each of three chamber tests and sent to each organization for counting. These three tests were selected to represent light, medium, and heavy densities of fiber deposits so as to provide a comparison over a range of conditions. The results of this counting are shown in Table 6.1.

It should be noted that there is much more variability in the counting procedure than these numbers suggest. For example, NASA counted sample 20812 four times and developed estimates ranging from 13800 to 19900 or a variation of ±23% from the mean in four observations. The sample variance of these observations furthermore is 1.9×10^7, suggesting a population variance of 2.5×10^7, close to the theoretical value in Table 5.1.

Apparently, all of the counting procedures have yielded the same results to within the theoretically estimated levels of error. It is further true that all methods provide accuracy to within the standards required for risk analysis of carbon fiber release.

7. CONCLUSIONS AND RECOMMENDATIONS

The preceding analysis has described in some detail the current methodologies for the counting of released carbon fibers and the associated sources of statistical errors. These errors have been quantified whenever possible as a means of evaluation and comparison of the techniques. The basic conclusion is that all procedures are sufficient for the estimation of released fibers and that the current Dugway Proving Ground procedure is to be preferred as it is much quicker and hence less expensive to implement. This procedure consists of placing a set of etched parallel lines on the samples, and counting fiber intersections with the lines to estimate the number of fibers. It has been modified recently by the replacement of wide lines with finely etched lines and by the introduction of 1-mm reference marks to aid in visual selection of fibers greater than 1 mm in length (shorter

TABLE 6.1

Comparison of Fiber Counts by Different Organizations

Test Number	Sampler Number	DPG	TRW	SSI	NASA
BT-171	13863	1206	662	1088	730.
(light deposit)	13875	1077	674	0	860
	13882	876	499	816	700
	13884	857	400	1474	910
BT-237	20803	25516	18341	40000	27600
(medium deposit)	20805	14227	15108	29100	19900
	20810	12718	13358	23850	19900
	20812	18782	17102	29800	16800
BT-230	19968	52071	48869	-	45800
(heavy deposit)	19970	47745	52016	-	40600
	19980	39151	33794	-	32000
	20013	36883	31475	-	29800

fibers are not counted). Table 7.1 summarizes the error estimates derived in this chapter. It should be pointed out that the generally accepted standard of accuracy in estimations of fiber release is to within a factor of 2 or 3. Consequently, any method providing this kind of accuracy, with no serious systematic tendencies to over or under estimate is sufficient for the analysis of the Dahlgren chamber tests. Selection among methods meeting this criterion should be based on considerations of cost and convenience of implementation.

It must be stressed, however, that the error estimates developed in this section are very conservative estimates. They are generally based on the concept of a 95% confidence interval, i.e., if the counting is repeated only 5% of the result will differ from the true number of fibers by more than the error limits specified. In practice, most measurements will be considerably more accurate. For example, at least 70% of the results will be in a band one half the width of the 95% interval. The estimates are also conservative in that they are based on a distribution independent method (Chebyshev Inequality) which specifies the largest possible confidence interval for any distribution having the calculated variance. In fact, the errors are likely to be well represented by a Gaussian distribution and a corresponding reduction of the confidence intervals would occur.

In summary, all procedures provide sufficient accuracy for the reduction of data from the Dahlgren chamber tests. The Dugway method as presently implemented represents the easiest procedure and is therefore preferred. The primary difficulty in the counting seems to arise from historically inconsistent application of the procedure which resulted in the introduction of fairly large (factor 3) nonrandom errors into the data. These errors are a function of the person counting the sample and of the density of fiber deposit on the sticky paper. This makes comparison of test results difficult. A partial correction for these errors is offered in Section 4.4.

TABLE 7.1

Summary of Estimated Errors

	DPG	TRW	NASA
Systematic	less than 15%	0	0
Random	75%	5%	75%
Total	less than 90%	5%	75%

APPENDIX A

THEORETICAL BASIS OF THE DUGWAY ALGORITHM

THEORETICAL BASIS OF THE DUGWAY ALGORITHM

A.1 The Buffon Needle Problem

The Dugway Proving Ground (DPG) counting procedure is based on the calculated probability that a needle dropped on a set of parallel lines will cross one of them. (Figure A.1)

This problem was solved by Buffon in 1733 who proved that if the needle length is less than the line spacing, the probability that a needle dropped on the parallel lines will cross one of them is given by:

$$P = \frac{2\ell}{\pi D} \tag{A.1}$$

where ℓ = fiber length

D = line spacing

This result may be demonstrated by the following argument. The location of the fiber relative to the grid lines may be specified by a normal distance x and an angle θ as shown in Figure A.1. Under the assumption that the fiber is equally likely to fall anywhere on the grid, the probability distributions for x and θ are as follows:

$$p(x) = \begin{cases} \frac{1}{D}, & 0 \leq x \leq D \\ 0 & \text{otherwise} \end{cases} \tag{A.2}$$

$$p(\theta) = \begin{cases} \frac{1}{\pi}, & 0 \leq \theta \leq \pi \\ 0 & \text{otherwise} \end{cases} \tag{A.3}$$

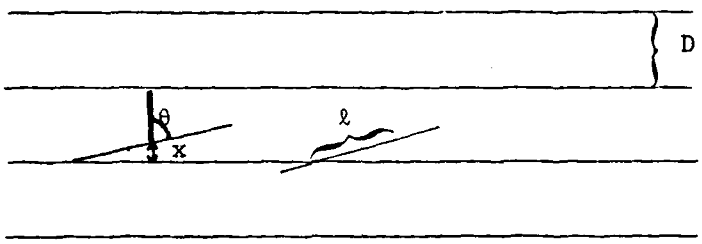
For each $x \leq \frac{D}{2}$, the fiber intersects the lower line if and only if $\theta \leq 2\cos^{-1}\frac{2x}{\ell}$ which implies (note that by symmetry the probability of intersecting the upper line is the same) that the intersection probability is given by:

$$P = 2\int_0^{\frac{\ell}{2}} \int_0^{2\cos^{-1}\frac{2x}{\ell}} \frac{1}{D\pi}\, d\theta dx$$

$$= \frac{2\ell}{\pi D} \tag{A.4}$$

FIGURE A.1

The Needle Intersection Problem



D = line spacing

ℓ = needle length

The Dugway algorithm was developed from this probability estimate (Solomon, et al.) by arguing that if a large number of needles or fibers are dropped on a set of parallel lines, the portion that will cross the lines is given approximately by this intersection probability. Thus, if N fibers are deposited, the number of fibers crossing the parallel lines should be approximately:

$$NP = \frac{2\ell N}{\pi D} \tag{A.5}$$

This suggests that the number of fibers deposited on a sticky paper may be estimated by covering the paper with parallel lines of known spacing, counting the number of fibers crossing the lines and estimating the total fiber count by rearranging Equation A.5 to read:

$$N = \frac{\pi D I}{2\ell} \tag{A.6}$$

where I is the number of intersections that were counted.

In the present application, the number of fibers deposited and, hence, the number of line crossings is still very large, so the Dugway algorithm introduces a further modification. Under the assumption that the fibers are uniformly distributed over the sheet, it should be possible to count line crossings over a smaller area, apply Equation A.6, and multiply the result by the ratio of the sticky paper area to the area counted. If A is the area of the sticky paper, A_s is the area counted and L is the length of line covering the area counted, then an estimate for the total number of fibers on the sheet is given by:

$$N = \frac{\pi D I}{2\ell} \cdot \frac{A}{A_s} = \frac{\pi D I}{2\ell} \cdot \frac{A}{LD} = \frac{\pi I A}{2\ell L} \tag{A.7}$$

which is the result used in the DPG fiber counting procedure.

A.2 Extensions of the Buffon Needle Theory

The analysis just described is insufficient to justify the DPG procedure since it assumes the fibers to be equal length and shorter than the line spacing. Arthur D. Little, Inc., has therefore extended this theory to collections of fibers of different lengths, including the case where fibers may be longer than the line spacing.

The Buffon needle theory as expressed above can be immediately extended to the case of fibers of various lengths but shorter than the line spacing if the average fiber length is used in Equation A.3. This is apparant if one notes that the average number of line crossings (I) will then be given by:

$$E\{I\} = \sum_{i=1}^{N} \frac{2\ell_i}{\pi D} = \frac{2}{\pi D} \sum_{i=1}^{N} \ell_i = \frac{2N}{\pi D} \bar{\ell} \qquad \text{(A.8)}$$

where ℓ_i = length of $i^{\underline{th}}$ fiber

N = total number of fibers

D = line spacing

I = number of line crossings

$\bar{\ell}$ = average fiber length

However, in the current problem, a significant number of fibers are longer than the line spacing because the spacing was selected to be close to the average fiber length. This requires further extension of the Buffon needle problem. Examination of the experimental data reveals that few fibers are longer than twice the average length, so it is sufficient to confine the analysis to fibers shorter than twice the line spacing. Under these assumptions and assuming that fiber is equally likely to fall anywhere on the sticky paper, it can be demonstrated (Appendix B) that the fiber intersection probabilities become:

Notation	Number of Intersections	Probability
$P_1^{(\ell)}$	1	$\frac{2\ell}{\pi D}$ if $0 \leq 2 \leq D$
		$\frac{2}{\pi}\left[2\cos^{-1}\frac{D}{\ell} - 2\sqrt{\frac{\ell^2}{D^2} - 1} + \frac{\ell}{D}\right]$ if $D \leq \ell \leq 2D$
$P_2^{(\ell)}$	2	0 if $0 \leq \ell \leq D$
		$\frac{2}{\pi}\left[\sqrt{\frac{\ell^2}{D^2}} - 1 - \cos^{-1}\frac{D}{\ell}\right]$ if $D \leq \ell \leq 2D$

where ℓ = fiber length

D = line spacing

It then follows that the average number of line crossings will be given by:

$$E\{I\} = \sum_{i=1}^{N} P_1(\ell_i) + \sum_{i=1}^{N} 2P_2(\ell_i) \qquad \text{(A.9)}$$

Thus, if there are N_1 fibers shorter than the line spacing, the following expression results:

$$E\{I\} = \sum_{i=1}^{N_i} \frac{2\ell_i}{\pi D} + \sum_{i=1}^{N} \frac{2}{\pi}\left[2\cos^{-1}\frac{D}{\ell_i} - 2\sqrt{\frac{\ell_i^2}{D^2} - 1} + \frac{\ell_i}{D}\right]$$

$$+ \sum_{i=N_1+1}^{N} 2 \bullet \frac{2}{\pi}\left[\sqrt{\frac{\ell_i^2}{D^2} - 1} - \cos^{-1}\frac{D}{\ell_i}\right]$$

$$= \sum_{i=1}^{N_1} \frac{2\ell_i}{\pi D} + \sum_{i=N_1}^{N} \frac{2}{\pi}\,\frac{\ell_i}{D} = \sum_{i=1}^{N} \frac{2\ell_i}{\pi D} = \frac{2N\bar{\ell}}{\pi D} \qquad \text{(A.10)}$$

Confirming the accuracy of the above procedure for distributions containing fibers longer than the line spacing.

APPENDIX B

SINGLE FIBER MULTIPLE INTERSECTION THEORY

SINGLE FIBER MULTIPLE INTERSECTION THEORY

The analysis in Appendix A assumes knowledge of the intersection probabilities for fibers longer than the line spacing along with the expected value and variance of the number of fiber intersections.

The following analysis has been developed to provide these values. The basis for these calculations are the following assumptions:

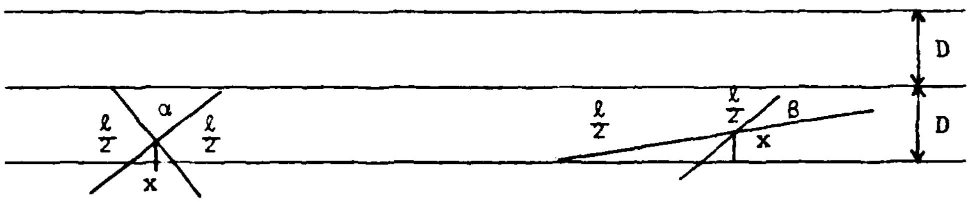
1. Each fiber's fall on the sheet is a statistically independent event.

2. The fiber's center of mass is equally likely to fall anywhere on the sheet.

3. The fiber is equally likely to lie at any angle relative to the parallel lines.

It then follows that the probabilities of observing a particular number of intersections is given by the fractional ranges (relative to π radians) of angles and (relative to line spacing) of fiber mid-point locations that will result in the number of intersections being considered. The following pages illustrate the geometric configuration leading to 1 or 2 intersections and the resulting probability calculations.

Calculation of probability of observing two intersections (P_{I2}) when $2D > \ell > D$



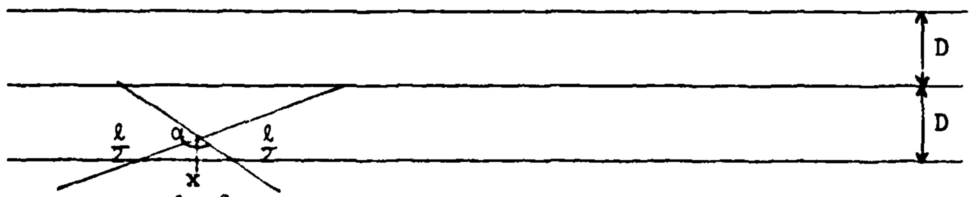
Since $\alpha = 2\cos^{-1}\frac{D-x}{\frac{\ell}{2}} = 2\cos^{-1}\frac{2(D-x)}{\ell}$, and since two intersections occur whenever the fiber touches the most distant grid line, it follows that

$$P_{I2} = 2\int_{D-\ell/2}^{D/2} \frac{2}{D\pi} \cos^{-1} \frac{2(D-x)}{\ell}\, dx$$

The above integral can be evaluated easily and results in

$$P_{I2} = \frac{2}{\pi}\left[\sqrt{\left(\frac{\ell}{D}\right)^2 - 1} - \cos^{-1}\left(\frac{D}{\ell}\right)\right]$$

Calculation of probability of one or more intersections (P_I) when $2LD > \ell > D$



Thus, $\alpha = 2\cos^{-1}\frac{2x}{\ell}$ represents the angular region over which one or more intersections will occur. From this it follows that:

$$P_I = 2\int_{o}^{D/2} \frac{2}{L\pi} \cos^{-1} \frac{2x}{\ell}\, dx$$

By evaluating the above integral we get

$$P_I = \frac{2}{\pi}\left[\cos^{-1}\frac{D}{\ell} - \sqrt{\frac{\ell^2}{L^2} - 1} + \frac{\ell}{D}\right]$$

<u>Calculation of probability of exactly 1 intersection (P_{I1}) when $2D > \ell > D$</u>

Since $P_{I1} = P_I - P_{I2}$, it follows from the above results that

$$P_{I1} = \frac{2}{\pi}\left[2\cos^{-1}\frac{D}{\ell} - 2\sqrt{\frac{\ell^2}{D^2} - 1} + \frac{\ell}{D}\right]$$

<u>Calculation of the expected number of intersections</u>

Under the assumption that each fiber's fall is a statistically independent event, the following provides an expression for the expected number of intersections.

$$E\left\{I\right\} = \sum_{i=1}^{N_1} \frac{2\ell_i}{\pi D} + \sum_{i=N_1+1}^{N} \frac{4}{\pi}\left[\sqrt{\frac{\ell_i^2}{D^2} - 1} - \cos^{-1}\frac{D}{\ell}\right]$$

$$+\sum_{i=N_1+1}^{N} \frac{2}{\pi}\left[2\cos^{-1}\frac{D}{\ell_i} - 2\sqrt{\frac{\ell_i^2}{D^2} - 1} + \frac{\ell_i}{D}\right]$$

where N = total number of fibers

and N_1 = number of fibers shorter than L

$$= \sum_{i=1}^{N} \frac{2}{\pi}\frac{\ell_i}{D} = \frac{2N\bar{\ell}}{\pi D}$$

where $\bar{\ell} = \frac{1}{N}\sum_{i=1}^{N} \ell_i$

Calculation of V(I), the variance of the number of intersections

Consider a variable x_i with following distributions:

x_i	$p_i \ (x_i)$
2	$\frac{2}{\pi}\left[\sqrt{\frac{\ell_i^2}{D^2} - 1} - \cos^{-1} \frac{D}{\ell_i}\right] \bullet \frac{A_s}{A}$
1	$\frac{2}{\pi}\left[- 2\sqrt{\frac{\ell_i^2}{D^2} - 1} + 2 \cos^{-1} \frac{D}{\ell_i} + \frac{\ell_i}{D}\right] \bullet \frac{A_s}{A}$
0	$1 - p_i(2) - p_i(1)$

It follows that

$$E\left\{x_i\right\} = \frac{2\ell_i}{\pi D} \quad \frac{A_s}{A}.$$

and

$$V\left\{x_i\right\} = E\left\{x_i^2\right\} - E^2\left\{x_i\right\}$$

$$= \frac{2}{\pi}\left[2\sqrt{\frac{\ell_i^2}{D^2} - 1} - 2\cos^{-1}\frac{D}{\ell_i} + \frac{\ell_i}{D}\right]\frac{A_s}{A} - \frac{4\ell_i^2}{\pi^2 D^2} \bullet \frac{A_s^2}{A^2}$$

This expression represents the variance of the number of intersections observed when a fiber of length ℓ_i falls on a set of parallel lines of spacing D, with $\ell_i > D$. If $\ell_i < D$, the variance becomes

$$V\left\{x_i\right\} = \frac{A_s}{A} \frac{2\ell_i}{\pi D} \left(1 - \frac{2\ell_i}{\pi D} \frac{A_s}{A}\right)$$

Since each fiber's fall on the sticky paper is considered to be a statistically independent event, the variance of the total number of intersections is given by the sum of the variances for each fiber or,

$$V\left\{ I \right\} = \sum_{i=1}^{N} \frac{A_s}{A} \frac{2\ell_i}{\pi D} \left(1 - \frac{A_s}{A} \frac{2\ell_i}{\pi D}\right) + \sum_{i=N_1+1}^{N} \frac{A_s}{A} \frac{4}{\pi} \left[\sqrt{\frac{\ell_i^2}{D^2} - 1} - \cos^{-1} \frac{D}{\ell_i} \right]$$

APPENDIX C

DESCRIPTION OF DISTRIBUTIONS USED IN THE ANALYSIS

DESCRIPTION OF DISTRIBUTIONS USED IN THE ANALYSIS

C.1 Exponential Distribution

This distribution has been used in some of the above calculations to permit rough estimation of the magnitudes of errors. Although other distributions, such as the incomplete Gamma, provide much better fits, the exponential is of a simple enough form to allow analytic calculation.

The density functions for the exponential distributions is given as follows:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } o \le x \le \infty \\ o & \text{otherwise} \end{cases}$$

From this, it follows that the cumulative distribution function is given by

$$F(x) = \int_{-\infty}^{x} f(x)\, dx = 1-e^{-\lambda x}$$

The mean and variance of this distribution are as follows:

$$\mu = \frac{1}{\lambda}, \quad \sigma^2 = \frac{1}{\lambda^2}$$

Table C.1 illustrates the fit which may be obtained for one of the tests.

C.2 Binomial Distribution

This distribution is used to model the results of a binary experiment with success probability a specified value p. It is described by the following probabilities:

$$\begin{aligned} f(x) &= \Pr \left\{ x \text{ successes during } n \text{ trials} \right\} \\ &= \binom{n}{x} p^x q^{(n-x)} \end{aligned}$$

TABLE C.1

Comparison of Exponential Distribution Function to Observed Data

FIBER LENGTH (mm)	Cumulative Fraction EXPONENTIAL	OBSERVED DISTRIBUTION	FRACTIONAL ERROR
2	0.27	0.28	-0.04
3	0.37	0.33	.11
4	0.46	0.42	.09
5	0.54	0.47	.13
6	0.60	0.54	.10
7	0.67	0.58	.13
8	0.71	0.65	.08
9	0.75	0.70	.07
10	0.79	0.76	.04
11	0.82	0.82	0.0
12	0.84	0.83	.01

$$= \binom{n}{x} p^x q^{(n-x)}$$

where $\binom{n}{x} = \dfrac{n!}{x!\,(n-x)!}$

The expected value and variance of the distribution are given by

$$\mu = n\,p$$

$$\sigma^2 = npq$$

DISTRIBUTION

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314 (12)

Library of Congress
Washington, DC 20540
ATTN: Gift and Exchange Division (4)

GIDEP Operations Office
Corona, CA 91720

Local:

E41
F56 (50)
X210 (2)